



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :  A01H 1/00, 1/04, C12N 15/00, 15/05, 15/09, 15/29, C12P 13/08, 21/00, 21/02		A1	(11) International Publication Number: <b>WO 97/07665</b>
			(43) International Publication Date: <b>6 March 1997 (06.03.97)</b>
(21) International Application Number: <b>PCT/US96/13538</b>		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: <b>22 August 1996 (22.08.96)</b>		Published <i>With international search report.</i>	
(30) Priority Data:  60/002,918 29 August 1995 (29.08.95) US 60,004,266 25 September 1995 (25.09.95) US			
(71) Applicant: UNIVERSITY OF HAWAII [US/US]; Suite 280, 2800 Woodlawn Drive, Honolulu, HI 96833 (US).			
(72) Inventors: SUN, Samuel, S., M.; The Chinese University of Hong Kong, Dept. of Biology, Shatin, New Territories (HK). XIONG, Liwen; University of Hawaii, Dept. of Plant Molecular Physiology, St. John #503, 3190 Maile Way, Honolulu, HI 96822 (US). LIU, Bolin; Institute of Botany, Academia Sinica, Xiangshan, Nanxincun 20, Beijing 100093 (CN). JING, Yuxiang; Institute of Botany, Academia Sinica, Xiangshan, Nanxincun 20, Beijing 100093 (CN).			
(74) Agents: ADLER, Reid, G. et al.; Morrison & Foerster L.L.P., 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).			
(54) Title: PRODUCTION OF TRANSGENIC PLANTS COMPRISING THE WINGED BEAN LYSINE- RICH PROTEIN			
(57) Abstract			
<p>A lysine-rich protein gene has been cloned from winged bean, <i>Psophocarpus tetragonolobus</i> and its DNA sequence and deduced protein sequence are disclosed. Methods for transforming plant cells are provided that mature into plants having elevated levels of lysine in their seeds.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

## PRODUCTION OF TRANSGENIC PLANTS COMPRISING THE WINGED BEAN LYSINE RICH PROTEIN

Field of the Invention

5 The present invention relates to the improvement of seed proteins. More particularly, the invention is directed to a lysine-rich protein isolated from winged bean (WBLRP), its associated nucleic acids and their uses, for example, as expressed in transgenic plants to improve their 10 nutritional qualities.

Background of the Invention

Plants are the primary source of all protein consumed by humans and livestock. In 1990, the world 15 protein supply was 150 million tons, of which 65% was obtained directly from plants and 35% through animal protein production. In comparison to meat, plant protein is much more economical to produce, store, and transport. However, when used as a source of dietary protein for human and 20 monogastric animals, most plant proteins are nutritionally incomplete due to their deficiency in several essential amino acids. In general, cereal (e.g., corn, wheat) proteins are low in lysine (Lys) and tryptophan (Trp), while legume (e.g., soybean) and most vegetable proteins are 25 deficient in the sulfur amino acids, methionine (Met) and cysteine (Cys). See, e.g., Sun and Larkins, in Transgenic Plants, vol. 1, King and Wu, Ed. at p. 317 (1992).

To overcome amino acid deficiencies, a mixed diet of cereal and legume, or supplementing plant proteins with 30 meat products or synthetic essential amino acids, typically is employed. U.S. growers annually spend over \$120 million

and \$70 million, respectively, on synthetic Met and Lys to supplement the corn and/or soy protein base of swine and poultry diets. In view of such economic factors, a more direct and cost-efficient means of supplying these amino acids in livestock feed would be highly desirable. See, e.g., Beach et al., in Biosynthesis and Molecular Regulation of Amino Acids in Plants, Singh et al., Eds., at p. 229 (1992).

Plant geneticists and breeders have made 10 significant efforts in the past to improve the quality of seed proteins and have developed, e.g., mutant high-lysine corn and barley. However, efforts in breeding legumes containing protein with increased levels of sulfur amino acids have not met with significant success.

15 Various genetic approaches to the selection or engineering of plant improvements include the selection of cultivars with high levels of desired amino acids, modification of amino acid biosynthesis, and modification of existing proteins. Unfortunately, there often are 20 undesirable traits associated with such modified plants, such as lower yields and greater susceptibility to pests and diseases, and these traits have prevented the agronomic utilization of the plants.

With respect to the selection of high seed protein 25 soybean cultivars, Serretti et al., Crop Sci. 34:207-209 (1994) compared the protein and amino acid content of several soybean lines. Similarly, Zarkadas et al., J. Agric Food Chem. 41:6161-623 (1993), compared the total protein amino acids and 4-hydroxyproline-rich glycoproteins of a 30 widely-grown soybean cultivar and a high-protein genotype. In one case, however, a high-lysine protein in barley was

found to be an inhibitor of chymotrypsin, arguably not a good candidate for nutritional supplementation because chymotrypsin inhibitors can interfere with digestion.

Similarly, U.S. Patent No. 5,367,110 (1995) to 5 Galili et al. describes transgenic plants that overexpress threonine and lysine. In such plants, a co-transformation process with chimeric genes introduces aspartate kinase activity as well as dihydridopicolinate synthase activity. This process is intended to enhance the biosynthesis of 10 aspartate-family amino acids, such as lysine, threonine, methionine and isoleucine, in particular the first two amino acids identified.

Alternatively, U.S. Patent No. 5,258,300 (1993) to Glassman et al. is directed to methods of inducing lysine 15 overproduction in plants by introducing the gene for dihydridopicolinic acid synthase (DHDPS) which is substantially resistant to feedback inhibition by endogenously-produced free L-lysine. Attached at the 5' terminus of this gene is a sequence encoding a chloroplast 20 transit peptide, which serves to localize the DHDPS in the chloroplasts of transformed cells, where it can act to enhance the biosynthesis of free L-lysine.

Other recent advances in agricultural biotechnology offer other molecular approaches to alter the 25 amino acid composition of seed proteins for improved nutritional quality. The transfer and expression of heterologous genes in plant tissues and plant seeds has been considered promising. It has been demonstrated that a seed protein gene can be transferred and expressed in transgenic 30 plants in a developmentally regulated manner, i.e., the protein product generally is synthesized, transported and

properly deposited in the seeds. Moreover, the feasibility of transforming a plant with an essential amino acid-rich protein gene also has been demonstrated.

For example, the research group of one of the 5 present inventors (Dr. Samuel Sun) isolated from Brazil nut a gene encoding a sulfur-rich protein (SRP), having 18% Met and 8% Cys. The SRP gene was subsequently (1989) transferred into tobacco plants to test if the expression of this transgene could enhance the Met content in the 10 transgenic seeds. Results showed that it is feasible to increase the Met content of the transgenic tobacco seeds (by up to 30%) through this molecular approach. See, e.g., Altenbach et al., Plant Molecular Biology 13:513-522 (1989). The SRP gene was acquired by Pioneer-Hibred International, 15 Inc., patented by Sun et al., and used for crop protein improvement.

However, what has been needed is method to transform plants in order to achieve overexpression of lysine without the problems and disadvantages associated 20 with prior art approaches. With respect to the expression of heterologous genes, the development of such methods will benefit from a better understanding of how gene transcription is regulated and the structure of seed storage proteins, the development of procedures for transforming and 25 regenerating important seed crops, and the identification and isolation of seed storage protein genes that encode proteins containing high levels of a specific essential amino acid or a combination of these amino acids.

Summary of the Invention

This invention relates to novel lysine rich proteins isolated from winged bean, and to its variants and 5 modifications.

In still other aspects, the invention provides an isolated nucleic acid molecule encoding a WBLRP and to its allelic variants, labeled or unlabeled, and a nucleic acid sequence that is complementary to, or hybridizes under 10 stringent conditions to, a nucleic acid sequence encoding a WBLRP.

Yet additional aspects of the present invention relates to transformed plant cells and intact plants derived from such cells that express elevated levels of lysine in 15 their seeds.

Brief Description of the Drawings

Fig. 1 shows the cDNA sequence of the 18 kD seed protein from winged bean.

20 Fig. 2 show the deduced amino acid sequence of the 18 kD seed protein from winged bean.

Modes of Carrying Out the Invention

We have identified and isolated a Lys-rich protein (WBLRP) from winged bean *Psophocarpus tetragonolobus* and a 5 cDNA encoding the WBLRP was cloned and sequenced (Fig. 1). The deduced amino acid sequence revealed that the 18 kD WBLRP contains 10.83 mol% Lys, representing the most abundant amino acid in this protein (Fig. 2). We have further demonstrated that this gene is expressed only in the 10 developing seeds, and therefore was a good candidate gene for the genetic engineering of seed proteins.

The WBLRP gene can be utilized, through transgenic technology, to enhance the Lys levels of plant proteins. It is particularly valuable in improving the nutritional 15 quality of cereal proteins, including those of corn, wheat, rice and barley, the world's major food crops, which are deficient in Lys. The economic and human nutritional values of this enhancement would be enormous.

As noted above, however, notwithstanding a strong 20 economic motivation to improve plants in this way, prior to our invention, no entirely useful products, without significant detrimental traits, have been reported. The WBLRP of the present invention also can be applied to enhance the Lys content in soybean for use as animal feed, 25 and there are recent studies on the economic implications of this enhancement. For example, the competitive product for Lys-rich soybean meal would be crystalline lysine-HCL, which costs about \$1.65/lb and contains 78% lysine. Therefore, a soybean meal contains 1% increased Lys would be worth about 30 \$55 more per ton. The estimated total additional value of increasing Lys by 1% in soybean as animal feed is \$851

million/year (Economic Implications of Modified Soybean Traits - Special Report 92, Iowa State University and Iowa Soybean Promotion Board & American Soybean Association cooperating, 1990).

5           In addition to the specific techniques described below, the practice of the present invention will employ conventional techniques of molecular biology, microbiology, recombinant DNA technology, and plant science, all of which are within the skill of the art. Such techniques are  
10 explained fully in the literature. See, e.g., Maniatis et al., Molecular Cloning: A Laboratory Manual (1982); DNA Cloning: Volumes I and II (D. N. Glover ed. 1985); Oligonucleotide Synthesis (M. J. Gait ed. 1984); Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1985);  
15 Transcription and Translation (B. D. Hames & S. J. Higgins, Eds. 1984); Animal Cell Culture (R. I. Freshney ed. 1986); Plant Cell Culture (R. A. Dixon ed. 1985); Propagation of Higher Plants Through Tissue Culture (K. W. Hughes et al. eds. 1978); Cell Culture and Somatic Cell Genetics of Plants  
20 (I. K. Vasil ed. 1984); Fraley et al. (1986) CRC Critical Reviews in Plant Sciences 4:1 (hereinafter Plant Sciences); Biotechnology in Agricultural Chemistry: ACS Symposium Series 334 (LeBaron et al. eds. 1987); and Genetic Engineering, Vol. 6 (Rigby ed., 1987).

25           In describing the present invention, the following terminology will be used in accordance with the definitions below.

          A "replicon" is any genetic element (e.g., plasmid, cosmid, chromosome, virus, etc.) that behaves as an  
30 autonomous unit of DNA replication in vivo; i.e., capable of replication under its own control within a cell.

A "vector" is a replicon, such as a plasmid, cosmid, or bacteriophage, to which another DNA segment may be attached so as to bring about replication of the attached segment, or to allow its introduction into a cellular host.

5 A "DNA molecule" is a species of nucleic acid, as defined below, and refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, or cytosine) in either single- or double-stranded form. The double-stranded form may, of course, be implied by the 10 context of the molecule's environment (e.g., a chromosome).

When in the double-stranded form, the molecule will usually be in its normal, double-stranded helix. The term "DNA molecule" is not limited to any particular tertiary form of DNA. Thus, the term includes double-stranded DNA found, 15 *inter alia*, in linear DNA molecules, viruses, plasmids, and chromosomes. When discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal conventions.

A DNA "coding sequence" is a DNA sequence which is 20 transcribed and translated into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by the start codon at the 5' terminus and a translation stop codon at the 3' terminus. Examples of 25 coding sequences include cDNA reverse transcribed from eukaryotic mRNA, genomic DNA sequences from eukaryotic cells, and synthetic DNA sequences.

A cell which has been "transformed" by an exogenous DNA sequence is a cell into which the exogenous 30 DNA has been introduced. The exogenous DNA may be integrated (covalently linked) to chromosomal DNA making up

the genome of the cell, or it may remain extrachromosomal. "Stably" integrated DNA sequences are those which are inherited through chromosome replication by daughter cells or organisms (accounting for loss by Mendelian segregation).

5 This stability is exhibited by the ability to establish cell lines or clones comprised of a population containing the exogenous DNA.

A "clone" or "cell line" is a population of cells descended from a single cell or common ancestor by mitosis 10 and is capable of stable growth in vitro for many generations.

A composition of a first type of matter (e.g., a DNA molecule containing a coding region) is "substantially free" of a second type of matter (e.g., DNA molecules which 15 do not contain a coding region) if the composition is comprised of less than about 10% (weight/weight) of the second type of matter relative to the sum of the first and second types of matter. Preferably, the composition contains less than about 5% of the second type of matter, 20 most preferably less than about 1%.

Two sequences, such as nucleic acid sequences or amino acid sequences, are "substantially homologous" when the sequences are a minimum of about 80% homologous over a selected region, preferably at least about 90% homologous, 25 and more preferably at least about 95% homologous. Identity or homology with respect to a common sequence is defined herein as the percentage of nucleic acids or amino acid residues in the candidate sequence that are identical with the WBLRP DNA or protein disclosed herein, after aligning 30 the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any

conservative substitutions as part of the sequence identity.

None of N-terminal, C-terminal or internal extensions, deletions, or insertions into the WBLRP sequence shall be construed as affecting homology.

5           WBLRP "nucleic acid" is defined as RNA or DNA that encodes a WBLRP, or is complementary to nucleic acid sequence encoding a WBLRP, or hybridizes to such nucleic acid and remains stably bound to it under stringent conditions, or encodes a polypeptide sharing at least 80%  
10 sequence identity, preferably at least 90%, and more preferably at least 95%, with the translated amino acid sequences shown in Fig. 2. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbone or including  
15 alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acid, however, is defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under stringent conditions, or is  
20 complementary to nucleic acid encoding a WBLRP or a pathogenesis-related protein as described below.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium titrate/0.1% NaDODS04  
25 at 50° C., or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42° C. Another  
30 example is use of 50% formamide, 5 x SSC (0.75M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1%

sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50  $\mu$ g/ml), 0.1% SDS, and 10% dextran sulfate at 42° C., with washes at 42° C. in 0.2 x SSC and 0.1% SDS.

5 "Isolated" nucleic acid will be nucleic acid that is identified and substantially free from other nucleic acid molecules, as defined above in the case of DNA. The nucleic acid may be labeled for probe purposes, using any label known and described in the art as useful in connection with 10 diagnostic assays.

Of particular interest is a WBLRP nucleic acid that encodes a full-length molecule, including but not necessarily the native signal sequence thereof. Nucleic acid encoding full-length protein may be obtained as 15 described above by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures to secure DNA that is complete at its 5' coding end. Such a clone is readily 20 identified by the presence of a start codon in a reading frame with the original sequence.

DNA encoding an amino acid sequence variant of a WBLRP is prepared as described below or by a variety of methods known in the art. These methods include, but are 25 not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of a 30 WBLRP.

//

Techniques for isolating and manipulating nucleic acids are disclosed for example by the following documents:

U.S. Patent No. 5,030,576, U.S. Patent No. 5,030,576 and International Patent Publications WO94/11504 and WO93/03162.

5 See, also, Sambrook, J. et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989, and Ausubel, F.M. et al. *Current Protocols in Molecular Biology*, Vol. 2, Wiley-Interscience, New York, 1987.

10 The claimed WBLRP protein and analog molecules that are the subject of this invention include molecules having the WBLRP amino acid sequence; fragments thereof having a consecutive sequence of at least 10, 15, 20, 25, 30 or 40 amino acid residues from a common WBLRP sequence; 15 amino acid sequence variants of a common WBLRP sequence wherein an amino acid residue has been inserted N- or C-terminal to, or within, the WBLRP sequence or its fragments as defined above; amino acid sequence variants of the common WBLRP sequence or its fragment as defined above which have 20 been substituted by another residue. WBLRP polypeptides include those containing predetermined mutations by, e.g., homologous recombination, site-directed or PCR mutagenesis, and WBLRP polypeptides of other plant animal species, and alleles or other naturally occurring variants of the WBLRP; 25 derivatives of the commonly known WBLRP or its fragments wherein the WBLRP or its fragments have been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such 30 as an enzyme or radioisotope); glycosylation variants of WBLRP (insertion of a glycosylation site or deletion of any

glycosylation site by deletion, insertion or substitution of appropriate amino acid); and soluble forms of WBLRP.

The novel nucleic acids and their associated proteins and peptides of the present invention are 5 preferably those which share the enhanced lysine levels of WBLRP and are capable of expression in transformed plants, but such nucleic acids and protein fragments and variants exclude any WBLRP polypeptide heretofore made public, including any known protein or polypeptide of any plant 10 species, which is otherwise anticipatory under 35 U.S.C. § 102 as well as polypeptides obvious over such known protein or polypeptides under 35 U.S.C. § 103.

In accordance with the present invention, a nucleic acid molecule is provided which contains, e.g., a 15 DNA coding region that encodes a WBLRP seed storage protein of *S. tetragonolobus*. Related DNA molecules can encode either a genomic or cDNA sequence, and either a complete presequence, or a mature sequence. DNA molecules according to the present invention can be comprised of a coding region 20 which is entirely homologous to native DNA; e.g., cDNA or genomic DNA. Alternatively, the coding region can be partially or completely synthetic in nature; e.g., comprised of codons different than that found in the native plant, yet encoding substantially the same protein. This may be 25 particularly preferred when it is intended to express the coding region in a heterologous host, thereby allowing for the selection of host-preferred codons. See, e.g., U.S. Patent No. 4,356,270; and EPO Pub. No. 46,039. In a preferred embodiment of the present invention, the DNA 30 sequences encoding WBLRP will be "flanked by DNA sequences heterologous" to the WBLRP coding sequences. This means

that the coding sequences flanked at both the 5' and 3' ends by DNA which is not found in the comparable flanking positions in the *S. tetragonolobus* wild-type genome.

Examples of DNA constructs falling within this definition 5 include plasmids carrying WBLRP-encoding sequences, linear DNA sequences containing non-*S. tetragonolobus* DNA in the 5' and 3' positions, or a *S. tetragonolobus* chromosome transformed with a WBLRP-encoding sequence that integrates at a position other than the location of the wild-type gene.

10 Particularly preferred flanking, heterologous sequences are those elements which make up the expression cassettes described below.

DNA molecules encoding WBLRP can be prepared synthetically by known methods of oligonucleotide synthesis.

15 Synthetic coding sequences can be prepared from overlapping oligonucleotides whose sequence contain codons for the amino acid sequence of the storage protein. Such oligonucleotides are prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) *Nature*

20 291:756; Nambiar et al. (1984) *Science* 323:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311; Oligonucleotide Synthesis (above).

WBLRP coding sequences can also be isolated employing methods of screening cDNA or genomic libraries.

25 While short probes can be prepared from the nucleic acid sequences disclosed herein, it is preferred to employ the entire coding sequences disclosed herein as a probe when screening libraries for other variants. Due to the homology between the variants, isolation of additional DNA sequences 30 encoding different isoproteins is within the skill of the art by employing the nucleic acid sequences disclosed herein

as probes. See, e.g., Molecular Cloning: A Laboratory Manual (above); DNA Cloning: Vol. I & II (above); Nucleic Acid Hybridization (above).

The synthesis or isolation of DNA molecules 5 containing the described coding sequences permits those of skill in the art to prepare compositions of DNA molecules encoding WBLRP substantially free of DNA molecules which do not encode the lysiner-rich protein. The compositions of DNA molecules containing coding sequence according to the 10 present invention can be comprised of linear coding sequences, or coding sequences in cloning vectors. Numerous cloning vectors are known to those skilled in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning 15 in the host cells which they transform include bacteriophage lambda (E. coli), pBR322 (E. coli), pACYC177 (E. coli), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-E. coli gram-negative bacteria), pHV14 (E. coli and Bacillus 20 subtilis), pBD9 (bacillus), pIJ61 (Streptomyces), pUC6 (Streptomyces), actinophage phi C31 (Streptomyces), YIp5 (yeast), YCp19 (yeast), and bovine papilloma virus (mammalian cells). See, generally, DNA Cloning: Volumes I & II (above); and Molecular Cloning: A Laboratory Manual 25 (above).

In accordance with the present invention, DNA constructs are provided in which the above-described coding sequences are placed under the control of transcription regulatory sequences so that the WBLRP storage protein is 30 expressed in a host cell. Such DNA constructs, referred to as "expression cassettes", are comprised of the storage

protein coding sequence under the control of regulatory sequences, such as a promoter, ribosome binding site (for bacterial expression) and a transcription termination sequence or polyadenylation signal. Other regulatory sequences include operators, enhancers, and the like. The cassette is usually flanked by convenient restriction sites.

The selection of the appropriate regulatory sequence and coding sequence, as well as their assembly for expression in a particular host, is within the skill of the art.

10 It is usually desirable to include within the expression cassette a selectable marker. A "selectable marker" gene encodes a "selectable phenotype"; i.e., a phenotype of a cell or organism which allows for the identification or selection of cells expressing the

15 selectable marker gene. Well-known marker genes are known in the art, including, but not limited to, the gene for chloramphenicol acetyltransferase (CAT), neomycin phosphotransferase (neo<5>), neomycin phosphotransferase II (npt-II), nopaline synthase (NOS), hygromycin

20 phosphotransferase, the glyphosate resistance gene (EPSP), dihydrofolate reductase (mtx<R>), hypoxanthine phosphoribosyltransferase (hpt), and thymidine kinase (tk). Cells transformed with sequences encoding these proteins are able to survive on media which would otherwise be toxic to

25 the cell. Other types of markers, such as beta-galactosidase (lacZ) cause cells transformed therewith to change color under certain conditions, thus allowing for visual selection. The presence of the selectable marker in the expression cassette allows for the determination of

30 whether a particular cell has been transformed stably by the expression cassette. Selectable markers can be included

which function in the ultimate host which will express the storage protein, as well as additional markers which will function in intermediate hosts in which the construction of the expression cassette occurs.

5           The expression cassette is constructed so that the coding sequence is located within the cassette with the appropriate control sequences, the positioning and orientation of the coding sequence with respect to the control sequences being such that the coding sequence can be  
10 transcribed under the control of the regulatory sequences (i.e., by RNA polymerase which attached to the DNA molecule at the control sequences) in a host cell transformed by the expression cassette.

15           It is possible to assemble the expression cassette prior to inserting it into a cloning vector as described above. Alternatively, an expression cassette can be constructed by cloning the coding sequence directly into an expression vector already containing the appropriate regulatory sequences and a restriction site downstream from  
20 the promoter.

          Construction of expression vectors or expression cassettes for use in transforming microbial hosts (e.g., bacteria or yeast) may be desired to produce single-cell protein by fermentation to supplement the nutritive value of  
25 other protein sources. A number of prokaryotic expression vectors are known in the art which could be adapted for this purpose See, e.g., U.S. Patent Nos. 4,440,859; 4,436,815; 4,431,740; 4,431,739; 4,428,941; 4,425,437; 4,418,149; 4,411,994; 4,366,246; and 4,342,832. See also British  
30 Patent Nos. 2,121,054; 2,008,123; 2,007,675; and European Patent Pub. No. 103,395. Yeast expression vectors are also

known in the art. See, e.g., U.S. Patent Nos. 4,446,235; 4,443,539; 4,430,428; 4,546,082. See also European Patent Pub. Nos. 103,409; 100,561; and 96,491.

A preferred class of heterologous hosts for the 5 expression cassettes containing coding regions for WBLRP are eukaryotic hosts, particularly the cells of higher plants (dicots, monocots). Particularly preferred among the higher plants are those of agronomic value having edible parts, such as leguminous plants, including, but not limited to, 10 *Glycine max* (soybean), *Medicago sativa* (alfalfa), *Psophocarpus tetragonolobus* (winged bean), and *Vigna aconitifolia* (moth bean). Other bean cultivars, such as *P. vulgaris*, can be useful hosts. Crops in general, such as maize, may also be useful hosts.

15 Expression cassettes intended for use in higher plants will employ regulatory sequences functional in such plants. For example, promoters can be selected from the group consisting of plant promoters, plant virus promoters, and T-DNA promoters (from both Ti and Ri plasmids).

20 Specific T-DNA promoters known in the art include the nopaline synthase (NOS) promoter, and the octopine synthase (OCS) promoter. These promoters are constitutive promoters.

See, e.g., Plant Sciences (above). Examples of plant virus promoters include the 19S and 35S transcript promoters from 25 cauliflower mosaic virus (CaMV). See, e.g., Id.; Koziel et al. (1984) J. Mol. Appl. Genet. 2:549. Numerous plant promoters have been shown to work in heterologous systems, including, but not limited to, the pea small subunit RUBP carboxylase (pSS) promoter, Morelli et al. (1985) Nature 30 (London) 315:200; Broglie et al. (1984) Science 224:838; Herrera-Estrella et al. (1984) Nature (London) 310:115;

Coruzzi et al. (1984) EMBO J. 3:1671; the soybean small subunit RUBP carboxylase (SbSS) promoter, Facciotti et al. (1985) Bio/Technology 3:241; the maize zein promoter, Matzke et al. (1984) EMBO J. 3: 1525; the wheat chlorophyll a/b 5 binding protein promoter, Lampa et al. (1985) Mol. Cell. Biol. 5:1370-1378; soybean 7S alpha, conglycinin promoter, Beachy et al. (1985) EMBO J. 4:3047; the soybean glycinin G2 promoter; soybean heat-shock promoter, EPO Publication No. 159,884; and the french bean phaseolin promoter, 10 Sengupta-Gopalan et al. (1985) Proc. Natl. Acad. Sci. USA 82:3320; Murai et al. (1984) Science 222:476. Depending upon the application, it may be desirable to select from among the available promoters those which are tissue specific (e.g., seed, leave, etc.) and/or regulated (e.g., 15 light-induced, temperature or heat-induced, developmentally regulated, etc.).

Particularly preferred promoters are those which allow for expression of WBLRP in seeds of heterologous plants. Examples of such promoters include, but are not 20 limited to, the phaseolin promoter and the soybean 7S alpha' conglycinin promoter. Expression cassettes intended for seed-specific expression can employ either heterologous promoters such as these, or the homologous promoter. In like manner, expression cassettes can employ heterologous 25 signal peptides in a coding sequence, particularly those associated with the heterologous promoter, or WBLRP's own signal peptide.

Various methods of transforming cells of higher plants with expression cassettes according to the present 30 invention are available to those skilled in the art. See generally, Lichtenstein et al. in *Genetic Engineering*: Vol.

6, pp. 104-182 (Rigby ed. 1987). Among the most popular of transformation methods are those based on transformation vectors constructed of T-DNA from Ti or Ri plasmids. It is particularly preferred to use binary T-DNA vectors. See, 5 e.g., Plant Sciences, above. T-DNA based vectors have been shown to transform dicots and monocots, including, but not limited to, legumes (e.g., soybean, peas, and alfalfa), cotton, rape, tomato, Liliaceae (e.g., asparagus), Amaryllidaceae, etc. See, e.g., Plant Sciences, above, and 10 references sited therein; Pacciotti et al. (1985) Bio/Technology 3:241; Byrne et al. (1987) Plant Cell, Tissue and Organ Culture 8:3; Sukhapinda et al. (1987) Plants Mol. Bio. 8:209-216; Lorz et al. (1984) Proceedings of EEC-Symp.: In Vitro Techniques-Propagation and Long-Term Storage; Lorz 15 et al. (1985) Mol. Gen. Genet. 199:178; and Potrykus et al. (1985) Mol. Gen. Genet. 199:183.

Other transformation methods are available to those skilled in the art. Viral transformation vectors are known, such as those based on CaMV. See, e.g., Brisson et 20 al. (1984) Nature (London) 310:511; Gronenborn et al. (1981) Nature (London) 294:773. Transposons can also be used to transform plants, particularly monocots. Maize transposons include, for example, Ac and Ds transposons, as well as the Mu transposon. See, e.g., Plant Sciences, above, at p. 28.

25 In the absence of a suitable transformation vector for desired higher plant host, such cells can be transformed by the direct uptake of DNA including expression cassettes of the present invention. The uptake of foreign DNA in plant cells using various chemical agents is known. See, 30 e.g., Davey et al. (1980) Plant Sci. Lett. 18:307; Draper et al. (1982) Plant Cell Physiol. 23:1; Krens et al. (1982)

Nature (London) 296:72; Hain et al. (1985) Mol. Gen. Genet. 199:161; Hooykaas-van Slogteren et al. (1984) Nature (London) 311:763; Hernalsteens et al. (1985) EMBO J. 3:3039. Higher plant protoplasts have also been transformed by 5 techniques of electroporation. See, e.g., Fromm et al (1986) Nature (London) 319:791; Fromm et al. (1985) Proc. Natl. Acad. Sci. USA 82:5824; Potter et al. (1984) Proc. Natl. Acad. Sci. USA 81:7161. Plant cells may also be transformed with foreign DNA by micro-injection using known 10 techniques, or by high-velocity metal particles coated with RNA or DNA. See, e.g., Kline et al. (1987) Nature (London) 327:70.

Once transformed plant cells have been produced, it is usually desirable to grow transformed cells into 15 callous tissue or a cell line in suspension. Techniques for regenerating callous and for producing cell lines are known in the art. See, e.g., Plant Cell Culture, above; Propagation of Higher Plants Through Tissue Culture, above. Transformed cells can also be induced to undergo 20 organogenesis. Thus, transformed cells can be maintained as undifferentiated tissues, or in organ culture. These plant cell cultures or organ cultures can be used, therefore, to produce WBLRP. See, e.g., Flores in Biotechnology and Agricultural Chemistry; ASC Symposium Series 334, p.66 25 (LeBaron et al. eds., 1987).

Another important application of the present invention is the production of plants transformed with the coding sequences for WBLRP. Techniques for regenerating plants from tissue culture, such as transformed protoplasts 30 or callous cell lines, are known in the art. See, e.g., Bingham et al. (1975) Crop Sci. 15:719-721; Kao et al.

(1980) Z. Pflanzenphysiol Bd. 96:135-141; Reisch et al.  
(1980) Plant Sci. Lett. 20:71-77; U.S. Pat. No. 4,548,901;  
Wright et al. (1987) Plant Cell Rpts. 6:83; Christianson et  
al. (1983) Science 222:632; Hammatt et al. (1987) Plant  
5 Science 48:129; Ghazi et al. (1986) Plant Cell Rpts. 5:452;  
Barwale et al. 1986) Planta 167:473; Newell et al. (1985)  
Plant Cell Tissue Organ Culture 4:145; Phillips et al.  
(1981) Plant Cell Tissue Organ Culture 1:123; Eapen et al.  
(1986) Theor. Appl. Genet. 72:384; Krishnamurthy et al.  
10 (1984) Plant Cell Rpts. 3:30; Shekhawat et al. (1983) Plant  
Sci. Lett. 32:43; Wilson et al. (1985) Plant Sci. 41:61;  
Venstesnaran et al. (1985) In Vitro 21(3,II): 36A. The  
selection of an appropriate method is within the skill of  
the art.  
15 Set forth below are specific examples of the  
present invention which are intended for illustrative  
purposes only. These examples are not intended to limit the  
present invention in any manner. The references cited in  
this application are incorporated by reference in their  
20 entireties.

Example 1

A. Identification of Seeds Rich in Lysine  
- The Winged Bean  
25 1. Seeds from 10 plant species of six families,  
as listed below, were selected for total protein extraction  
and analysis.

30 *Corylus spp.* (hazel nut)  
*Brassica spp.* Chinese green mustard)  
*Carya cathayensis* (mountain walnut)  
*Juglans regia* (English walnut)

*Vigna radiata* (mung bean), 3 varieties  
*Psophocarpus tetragonolobus* (winged bean)  
*Sesame indicum* (sesame)  
*Pinus tabulaeformis* (Chinese pine)

5

2. The amino acid composition of the 10 protein samples was determined by an automated amino acid analyzer.
3. From the amino acid profiles of the 10 protein samples, winged bean was identified as having the 10 highest lysine content, 7.9% (g/100g).

B. Identification of Winged Bean Polypeptide(s)  
Rich in Lysine - the 18-KD Protein

1. The polypeptides in the total seed protein extract of winged bean were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) into different size species.
2. Twenty one individual polypeptides differing in molecular weight were excised and eluted from the SDS-PAGE gel and their amino acid compositions determined.
3. From the amino acid profiles of the 21 polypeptide species, an 18-KD polypeptide was identified to contain the highest amount of lysine, 10.17% (mg/100 mg), and named winged bean lysine-rich protein (WBLRP).

25

C. Determination of the N-Terminal  
Amino Acid Sequence of WBLRP

1. The N-terminal amino acid sequence of WBLRP was determined. The sequence of the first 8 amino acids as determined and confirmed is as shown below:

NH<sub>2</sub> - Gly-Val-Phe-Thr-Tyr-Glu-Asp-Glu - COOH

### Example 2

## Cloning the cDNA Encoding the Winged Bean 18-kD LRP

The cloning the cDNA encoding the WBLRP was  
5 achieved with the following approach: verify the presence  
of the mRNA encoding the WBLRP in the developing winged bean  
seeds; amplify and clone the mRNA encoding the WBLRP by PCR  
technique; construct a cDNA library for the developing  
winged bean seeds; screen the cDNA library using the PCR  
10 generated WBLRP cDNA as a probe; and sequence and further  
characterize the cDNAs encoding the WBLRP.

A. Identification of mRNA Encoding the 18-KD WBLRP in Developing Winged Bean Seeds

15 1. On the basis of the N-terminal sequence obtained for the (I.C2), a 17-mer oligonucleotide mixed probe, WB-NT1, with a degeneracy of 1/64 and complementary to the mRNA, was synthesized:

25 Corresponding C - Glu - Asp - Glu - Tyr - Thr - Phe -  
 WBLRP N- 8 7 6 5 4 3  
 terminal sequence:

30           2. Total RNA was extracted from maturing winged  
bean cotyledons (Hall, et al., 1978. Proc. Natl. Acad. Sci.  
USA 75:3196-3200) and separated by formaldehyde gel  
electrophoresis.

34

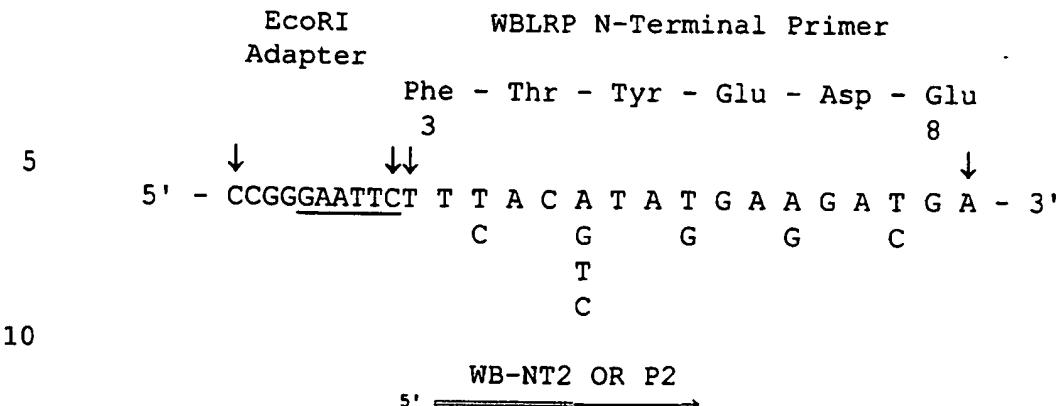
3. The winged bean RNA was transferred from the gel to nitrocellulose filter (northern blotting) and hybridized to the <sup>32</sup>P labeled 17-mer WB-NT1 probe.

4. A mRNA species with molecular size of 700 NT 5 was identified for the WBLRP, indicating that the RNA prepared from the developing winged bean cotyledons contains the mRNA for the WBLRP, and thus justifying the use of this RNA for cloning the LRP gene.

10 B. Amplification of the WBLRP mRNA and Cloning the Amplified Product

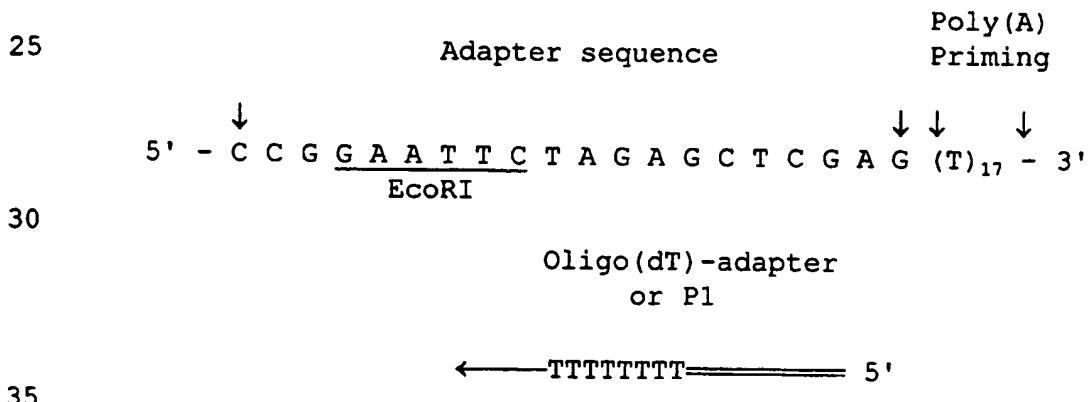
1. To amplify and clone the cDNAs encoding the WBLRP using polymerase chain reaction (PCR) technique, three oligonucleotide primers were synthesized, a WBLRP gene 15 specific primer, WB-NT2 or P2, and two universal primers for priming poly(A) RNA, Oligo(dT)-adapter or P1, and Adapter or P3.

20 This 26-mer oligonucleotide primer consists of a mixture of 17 nucleotides (degeneracy, 1/64) that are synthesized based on the N-terminal amino acid sequence of WBLRP, NH<sub>2</sub>-Phe-thr-Tyr-Glu-Asp-Glu (the 3rd to 8th amino acids) and a 10-nucleotide region contains an EcoRI 25 restriction enzyme site (CCGGGAATTC) to facilitate the cloning of the PCR products. P2 is the 5' PCR primer for amplifying the WBLRP mRNA.



15 - Primer Oligo(dT)-Adaptor or Pla:

This 37-mer primer consists of 17 (dT) nucleotides for priming the poly (A) tail of any mRNA and a 20-nucleotide region that contains multiple restriction sites including EcoRI to facilitate the cloning of PCR products. P1 is the 3' primer for transcribing the mRNA (to produce the first strand cDNA).

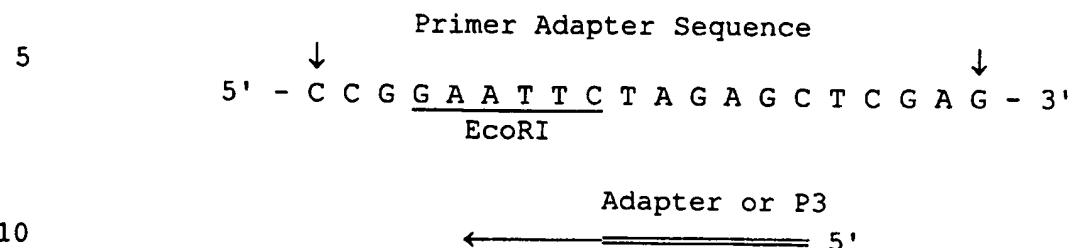


#### - Primer Adapter or P3:

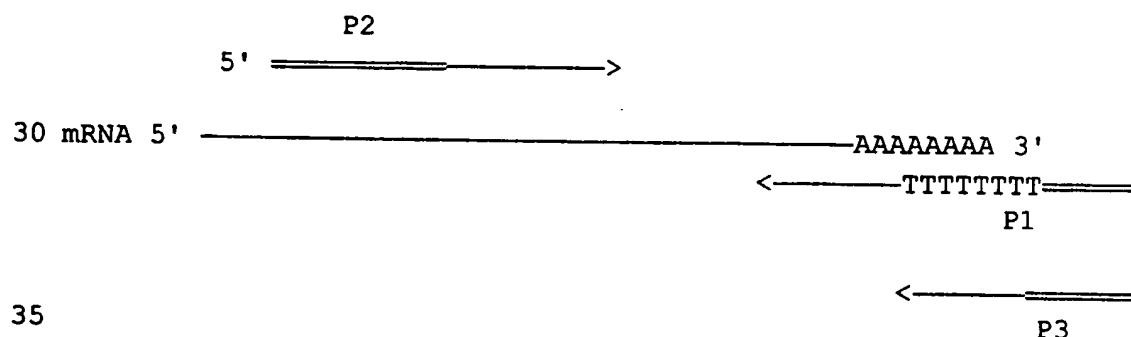
This 20-mer primer contains multiple restriction 40 sites including EcoRI. It is identical to the adapter

24

sequence of P1 and is used to prime the 2nd strand cDNA for continuous synthesis of the first strand cDNA.



2. The mRNA encoding the WBLRP was amplified by the PCR technique, RACE (Rapid Amplification of cDNA Ends) (Frohman, et. al., 1993. Methods in Enzymology 218:340-356).  
15 In this technique, all the mRNAs of the developing winged bean cotyledons were first reverse transcribed to generate a mixture of complementary DNAs (first strand cDNAs) using the P1 primer (Oligo(dT)-adapter) and reverse transcriptase.  
20 The first strand WBLRP cDNA in the pool of mixed cDNAs was then selectively transcribed (into 2nd strand cDNA) using the P2 (WB-NT2) gene specific primer by Taq polymerase. Primer P3 (Adapter) was then used to multiply the first strand LRP cDNA. A total of 30 cycle of amplification was  
25 carried out.



3. After amplification, the PCR products were separated by (2% low melting) agarose gel electrophoresis. Two major DNA bands, 700bp and 400bp, were observed. The 700-bp and 400-bp DNA fragments were excised and purified 5 from the gel (by Promega's Wizard™ PCR Preps-DNA Purification System). After digestion with ECoRI and purification through Millipore's Ultrafree-MC column, the 700-bp and 400-bp EcoRI DNA fragments were inserted into the pBluescript KS+ vector (stratagene) and transformed into *E.* 10 *coli* DH5 $\alpha$  competent cells.

4. Colonies with the vector and insert DNA were identified. The nucleotide sequences of the insert DNAs were determined by the dideoxy sequencing procedure (Hanger et al., 1977. Proc. Natl. Acad. Sci., USA 74:5463-5467) 15 using the T3 and T7 primers and the Sequence of United States Biochemical.

5. Comparison of the cloned nucleotide sequences with the known N-terminal WBBRP sequence revealed that the 700-bp PCR cDNA clones encode the WBLRP but not the 400-bp 20 DNA. The entire sequence of 742 bp (excluding the poly A tail) of clone WB7-28 was determined. The N-terminal sequence of the deduced protein sequence matches with that of the 18-KD WBLRP. The deduced amino acid sequence contains 155 residues (missing two because the primer starts 25 from the 3rd amino acid), of which 17 are lysine. This cDNA thus encodes a protein rich in Lys (10.83 mol% Lys).

C. Cloning the cDNA for WBLRP

The strategy of cloning a cDNA encoding the WBLRP is to first construct a cDNA library for the developing 5 winged bean cotyledons and then screen the library, using the PCR-cloned WBLRP cDNA as a probe, to identify a full-length cDNA encoding the WBLRP.

1. Poly(A) RNA was isolated from the total winged bean RNA by the PolyATtract mRNA Isolation System of 10 Promega.

2. A cDNA library was constructed by the ZAP cDNA Synthesis and Cloning System of Stratagene using 5 $\mu$ g of winged bean poly(A) RNA. A total of  $9 \times 10^6$  recombinant clones was obtained for the primary cDNA library and  $1.8 \times 15 \times 10^9$  recombinant clones for the amplified library.

3. The PCR-cloned WBLRP cDNA, WB7-28, was labeled with  $^{32}$ P by the method of random priming and used as a probe to screen the winged bean cDNA library. Positive clones were identified and their sequences determined.

20 4. The complete nucleotide sequence of cDNA clone WB-3 was determined. The total length of WB-3 is 782 bp excluding the poly A tail. It contains a 5' untranslated region of 68 nucleotides, a coding region of 474 nucleotides, and a 3' untranslated region of 240 nucleotides.

25 The coding region encodes a protein of 157 amino acids with a molecular weight of 19,494 daltons (thus 1,494 daltons more than the 18-KD estimated by SDS-PAGE). There are 17 residues of Lys, representing 10.83 mol% of the total amino acids.

Example 3Further Characterization of the WBLRP GeneA. Sequence Homology with the  
5 Pathogenesis-Related Proteins

A search of the sequence data in the GenBank revealed that the WBLRP share high sequence homology (50-87%) with those of the many pathogenesis-related (PR) proteins. PR proteins accumulate as a part of a defense 10 response system in many plants exposed to viroids, viruses, fungi, bacteria and other stresses (Linhorst, H.J.M., 1991. Critical Rev. Plant Sci., 10:123-150; Fristensky, B. et al., 1988. Plant Mol. Biol. 11:713-715; and Crowell, D.N. et al., 1992. Plant Mol. Biol. 18:459-466). Upon calculating the 15 lysine content of the available sequences of PR proteins, we found that these proteins, like WBLRP, are rich in lysine. However, in the literature on PR proteins, there was no noticeable attention paid to the high lysine content of these proteins, nor any recognition of the potential use of 20 these proteins for nutritional quality improvement. This may be due to the scientists working in the PR protein area center their interest only in the pathogenicity aspects of these proteins, or the involvement of the PR proteins in plant diseases renders these proteins unattractive for 25 potential use in plant nutrition improvement. Thus, the literature might reveal the utility of the PR proteins for disease resistance and not for plant nutrition improvements.

B. WBLRP: Special Biological Characteristics  
Favoring its use for Plant Nutrition Improvement

We have carried out further experiments to 5 investigate the molecular biology of the WBLRP. Unlike most PR proteins that are synthesized in response to pathogen infection and wounding, chemical and other stresses, we found that the WBLRP is a relatively abundant protein (2% as total protein) in the seeds of winged bean. Northern 10 hybridization experiments revealed that the expression of the WBLRP gene is organ specific, i.e. the WBLRP gene only expressed in the seed but not in the leaf, stem, root, or flower. The expression of the WBLRP gene is also developmentally regulated. Specifically, the expression of 15 the WBLRP in the developing seeds starts 20 days after flowering (DAF), peaks at 35 DAF, and declines 45 DAF. Thus the WBLRP gene is under tight control for spatial and temporal expression in the normally developing winged bean seeds; in this respect, its expression is not related to 20 pathogen activity or wounding. The synthesis and accumulation of the lysine-rich protein in abundance in the seeds makes the WBLRP gene a suitable target for plant (especially seed) nutrition improvement.

25

Example 4

The Construction of Transgenic Plants

Plants transformed with a vector comprising the WBLRP protein are prepared as follows. To illustrate, transformation of rapeseed for enhanced lysine content will 30 be described.

A chimeric gene containing the WBLRP coding sequence (cDNA) under the regulation of the French bean

seed-specific phaseolin promoter and terminator sequences will be constructed. This chimeric gene is then inserted into the *Agrobacterium* transformation vector, pBI121. The vector harboring the WBLRP chimeric gene will then be 5 transferred into *A. tumefaciens* LBA4404, and used to transform the rapeseed explants.

The transformation and regeneration of several commercial cultivars of rapeseed plants are well documented and established. To summarize, there are three major 10 procedures for rapeseed transformation: 1. stem segments as explants via *A. rhizogenes* (Guerche et al. *Mol. Gen. Genet.* 202:382-386, 1987) or *A. tumefaciens* (Pua et al. *Bio/Tech.* 5:815-817, 1987); 2. Hypocotyls as explants via *A. tumefaciens* (De Block et al., *Plant Physiol.* 91:694-701, 15 1989) or *A. rhizogenes*, Damgaard & Rasmussen, *Plant Mol. Biol.* 17:1-8, 1991); and 3. Protoplasts as starting materials via electroporation (Guerche et al., *Plant Sci.* 52:111-116, 1987). Using these procedures, transgenic rapeseed plants with normal phenotype and full fertility can 20 be generated in about four months.

The procedure using *A. tumefaciens* to transform the hypocotyl segments of rapeseed will be followed. In the present example, the *A. tumefaciens* LBA4404 carrying the WBLRP chimeric gene will be used to transform the rapeseed 25 cultivar Westar, a spring cultivar, and the transgenic plants will be recovered, as detailed by Altenbach et al., *Plant Mol. Biol.* 18:235-245.

The integration of the WBLRP chimeric gene into the rapeseed genome will be detected by Southern blot 30 hybridization using the WBLRP cDNA as a probe. The expression of the transgene in the transgenic seeds will be

determined by: 1) Northern blot hybridization, using WBLRP cDNA as a probe, for the WBLRP transcripts (mRNA), and 2) Western blot/immunodetection, using antibody specific for the LRP, for the WBLRP protein. To quantitate the lysine 5 content in the transgenic seeds, total protein will be extracted from the seeds and its amino acid composition determined. Comparison of the lysine contents between the transgenic and control (untransformed) seeds should reveal the enhancement of the lysine content in the transgenic 10 seeds.

Based on the foregoing disclosure, variations and modifications on the foregoing embodiments are within the ability of one of ordinary skill in the art, and such variations and modifications do not depart from the scope of 15 the present invention as described in the following claims.

Claims

1. A composition comprising DNA molecules  
5 containing a coding sequence encoding WBLRP or an allelic  
variant thereof, said composition being substantially free  
of DNA molecules that do not contain said coding sequence.

2. A composition according to claim 1 wherein  
said DNA molecules are replicons.

10

3. A DNA molecule containing a coding sequence  
encoding WBLRP or an allelic variant thereof wherein said  
coding sequence is flanked by heterologous sequences which  
are T-DNA.

15

4. A DNA molecule containing a coding sequence  
encoding WBLRP or an allelic variant thereof wherein said  
coding sequence is flanked by heterologous sequences which  
are plant virus nucleic acid sequences.

20

5. A plant cell comprising a DNA sequence  
according to claim 3.

6. A plant cell comprising a DNA sequence  
25 according to claim 4.

7. A recombinant DNA sequence containing an  
expression cassette, said cassette containing a coding  
sequence encoding WBLRP or an allelic variant thereof, said  
30 coding sequence being under the transcriptional and/or

translational control of regulatory sequences which are heterologous to said WBLRP-encoding DNA.

8. The recombinant DNA sequence of claim 7  
5 wherein the control sequences of the expression cassette include a promoter selected from the group consisting of the nopaline synthase promoter, the octopine synthase promoter, the pea small subunit RUBP carboxylase promoter, and the soybean small subunit RUBP carboxylase promoter, the maize 10 zein promoter, the wheat chlorophyll A/B binding protein promoter, soybean 7S-alpha'-conglycinin promoter, the soybean glycinin G2 promoter, soybean heat shock promoter, and the french bean phaseolin promoter.

15 9. A plant cell comprising the recombinant DNA sequence of claim 7.

10. A plant cell comprising the recombinant DNA sequence of claim 9.

20

11. A plant cell which has been transformed with the recombinant DNA sequence of claim 7.

12. A plant cell which has been transformed with 25 the recombinant DNA sequence of claim 8.

13. A plant comprising the recombinant DNA sequence of claim 7.

30 14. A plant comprising the recombinant DNA sequence of claim 8.

15. A plant regenerated from the cells of claim 9  
or the progeny of said plant.

5 16. A plant regenerated from the cells of claim  
10 or the progeny of said plant.

17. A method of increasing the level of free  
L-lysine in a plant comprising the steps of:

10 transforming a WBLRP-encoding gene into a  
susceptible plant cell;  
growing a plant from said cell; and  
expressing said WBLRP-encoding gene in said plant.

15 18. The method of claim 17 wherein a plasmid  
comprising said WBLRP-encoding gene is introduced into said  
cells.

19. The method of claim 17 wherein said plasmid  
20 is a bacterial cloning vector.

20. The method of claim 19 wherein said plasmid  
is an *E. coli* cloning vector.

25 21. The method of claim 17 wherein the WBLRP-  
encoding gene is introduced into said plant cells by the  
techniques of electroporation, microinjection,  
microparticles, or liposomal encapsulation.

22. The method of claim 19 wherein the WBLRP-encoding gene is introduced into the plant cells by Agrobacterium-mediated transformation.

5 23. The composition of claim 1 wherein said coding sequence has the sequence of Fig. 1.

1 / 2

AATTGGCAC GAGCGGCACG AGGCAAATC TCTTCTCAA ATCAACGTTA  
TTTTCCATT TCTATTTGTA ATTTGATCAT TATGGGTGTT TTCACATATG  
AGGATGAAAC CACTCACCA GTGGCTCCTG CTATCCTTA CAAAGCAATA  
GTTAAAGATG CTGATAACAT CTTTCAAAG GCTGTTGATT CCTTTAAGAG  
TGTTGAAATT GTTGAGGGAA ATGGTGGTCC TGGAACCATC AAGAAGATCT  
CTTTGTTGA GGATGGGAA AGCAAGTTG TGTTGCACAA GATTGAGTCA  
ATTGATGAGG CTAATTTGGG ATACAGCTAC AGCATAGTTG GTGGTGCTGC  
TTTGCCAGAC ACAGTGGAGA AGATTACATT TGAGTCCAAA TTGAGTGCTG  
GACCTTCTGG AGGCTCTGTT GGGAAACTCA CTGTGAAATA CCAAACCAAA  
GGAGATGCTG AGCCAATGA AGAGGAACTC AAAGTTGGCC CATCCTGAAT  
ACAATTGATA TCATATCAAC CATAATCCAA TTTCCAATTCA AACTTAATCA  
CACTCAGATG ATTATCATTAT TATGATAACCC TATGGTCTCA AATTAAATGA  
ATCATATATG <sup>\*\*</sup> ATATCAGTTT ATTAAGGTTT TTCAGGAGAG TTTTGTGTT  
TTGCAATGTT GTCCTGTGTG CTCAAGTTGG CTTTCAGTGT AATAATAAGT  
CCATGTTCAT TGGATTGAAG TATTTATAA ACTGAAGTTT GATGAAAAAA  
AAAAAAAAAA

**FIG. 1**

2 / 2

Met-Gly-Val-Phe-Thr-Tyr-Glu-Asp-Glu-Thr-Thr-Ser-Pro-Val-Ala  
Pro-Ala-Ile-Leu-Tyr-Lys-Ala-Ile-Val-Lys-Asp-Ala-Asp-Asn-Ile  
Phe-Pro-Lys-Ala-Val-Asp-Ser-Phe-Lys-Ser-Val-Glu-Ile-Val-Glu  
Gly-Asn-Gly-Gly-Pro-Pro-The-Ile-Lys-Lys-Ile-Ser-Phe-Val-Glu  
Asp-Gly-Glu-Ser-Lys-Phe-Val-Leu-His-Lys-Ile-Glu-Ser-Ile-Asp  
Glu-Ala-Asn-Leu-Gly-Tyr-Ser-Tyr-Ser-Ile-Val-Gly-Gly-Ala-Ala  
Leu-Pro-Asp-Thr-Val-Glu-Lys-Ile-Thr-Phe-Glu-Ser-Lys-Leu-Ser  
Ala-Gly-Pro-Ser-Gly-Gly-Ser-Val-Gly-Lys-Leu-Thr-Val-Lys-Thr  
Gln-Thr-Lys-Gly-Asp-Ala-Glu-Pro-Asn-Glu-Glu-Leu-Lys-Val  
Gly-Lys-Ala-Lys-Glu-Asp-Ala-Leu-Phe-Lys-Ala-Val-Glu-Ala-Tyr  
Leu-Leu-Ala-His-Pro-Glu-Tyr-Asn

## FIG. 2

## INTERNATIONAL SEARCH REPORT

International application No. PCT/US96/13538
---

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.

US CL : 800/205; 435/69.1, 172.3, 240.4, 320.1; 536/23.6, 24.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 800/205; 435/69.1, 172.3, 240.4, 320.1; 536/23.6, 24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, BIOSIS, CABAB, CAPLUS, MEDLINE

search terms: winged bean, lysine, protein, Psophocarpus tetragonolobus

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,367,110 A (GALILI et al) 22 November 1994, columns 1-14.	1-23
Y	US 5,258,300 A (GLASSMAN et al) 02 November 1993, columns 1-10.	1-23
Y	THOMPSON et al. Winged Bean: Unexploited Tropical Food Crop. Hortscience. June 1980, Vol. 15, No. 3, pages 233-238, especially pages 234-236.	1-23

Further documents are listed in the continuation of Box C.

See patent family annex.

•	Special categories of cited documents:	T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
•A	document defining the general state of the art which is not considered to be of particular relevance		
•E	earlier document published on or after the international filing date	X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
•L	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	Y	when the document is taken alone
•O	document referring to an oral disclosure, use, exhibition or other means	&	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
•P	document published prior to the international filing date but later than the priority date claimed		document member of the same patent family

Date of the actual completion of the international search

07 NOVEMBER 1996

Date of mailing of the international search report

20 NOV 1996

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231  
Facsimile No. (703) 305-3230

Authorized officer  
THOMAS HAAS  
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US96/13538

A. CLASSIFICATION OF SUBJECT MATTER:  
IPC (6):

A01H 1/00, 1/04; C12N 15/00, 15/05, 15/09, 15/29; C12P 13/08, 21/00, 21/02